LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-PROC-422574

Approximation Properties of Coarse Spaces by Algebraic Multigrid

P. S. Vassilevski

January 19, 2010

Fourth Annual Meeting of the Bulgarian Section of SIAM
Sofia, Bulgaria
December 21, 2009 through December 22, 2009

Disclaimer

Approximation Properties of Coarse Spaces by Algebraic Multigrid

Panayot S. Vassilevski *

January 14, 2010

Abstract

This note provides an application of the "window"-based spectral AMG method (cf. [FVZ05] or [Va08]) for proving energy error estimates desired in providing coarse (upscaled) discretization of fairly general classes of PDEs. Computationally more efficient versions of the original window-based spectral AMG, as well as a new method utilizing local element matrices, are outlined summarizing the results from a forthcoming report.

1 The strong approximation property

We are given a s.p.d. $n \times n$ sparse matrix A and let $P: \mathbf{R}^{n_c} \mapsto \mathbf{R}^n$, $n_c < n$, be a given (rectangular) interpolation matrix.
We are interested in the following *strong approximation property*:
For any fine-grid vector $\mathbf{u} \in \mathbf{R}^n$ there is a coarse interpolant $P\mathbf{u}_c$ such that

$$\|A\|\|\mathbf{u} - P\mathbf{u}_c\|_A^2 \leq C_A \, \|A\mathbf{u}\|^2. \tag{1.1}$$

If the problem of our main interest

$$A\mathbf{u} = \mathbf{f},$$

comes from a finite element discretization of a PDE on a domain $\Omega \subset \mathbf{R}^d$ ($d = 2$ or 3), then $\mathbf{f} = (f_i)$ comes from a given r.h.s. function $f(\mathbf{x}) \in L_2(\Omega)$, where the entries f_i are computed as the following integral moments

$$f_i = (f, \, \varphi_i) \equiv \int_\Omega f(\mathbf{x})\varphi_i \, d\mathbf{x}.$$

Above, φ_i runs over a basis of the fine–grid finite element space V_h associated with a triangulation of Ω with characteristic fine-grid mesh size h. For a nodal (Lagrangian)

*This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

basis, the index "i" runs over the set of fine degrees of freedom $x_i \in \mathcal{N}_h$. The unknown $\mathbf{u}$ stands for the coefficient vector of the finite element, Galerkin, approximation u_h to the solution of the underlined PDE posed variationally, i.e., $u_h \in V_h$ solves the discretized PDE in a variational form associated with a given bilinear form $a(.,\ .)$, stated as follows

$$a(u_h,\ \varphi) = (f,\ \varphi) \text{ for all } \varphi \in V_h.$$

As an example, we consider a second order self–adjoint elliptic bilinear form $a(u,\ \varphi) = \int\limits_{\Omega} k(\mathbf{x})\ \nabla u \cdot \nabla \varphi\ d\mathbf{x}$ for $u,\ \varphi \in H_0^1(\Omega)$ and a given positive coefficient function $k = k(\mathbf{x}),\ \mathbf{x} \in \Omega$, the given polygonal/polyhedral domain in $\mathbf{R}^d$, $d = 2$ or 3. Using a standard piecewise linear conforming finite element space V_h on a quasiuniform triangulation $\mathcal{T}_h$, as it is well-known, the stiffness matrix $A = (a(\varphi_j,\ \varphi_i))$ computed from a nodal Lagrangian basis $\{\varphi_i\}_{\mathbf{x}_i \in \mathcal{N}_h}$ of V_h satisfies

$$\|A\| \simeq h^{d-2}. \tag{1.2}$$

The equivalence constants above generally depend on the variation $\frac{\max\limits_{\mathbf{x}\in\Omega} k(\mathbf{x})}{\min\limits_{\mathbf{x}\in\Omega} k(\mathbf{x})}$ but are mesh-independent.

Assume, that we have come up with a coarse space $V_H \subset V_h$ such that the coefficient vectors of functions in V_H viewed as elements of V_h can be represented as the range of an interpolation mapping P. We can define respective coarse basis functions by forming $P\mathbf{e}_{i_c}$ for each coarse coordinate vector $\mathbf{e}_{i_c} \in \mathbf{R}^{n_c}$ that has a single nonzero entry at the i_cth position. Then, consider the fine–grid function $\phi_{i_c}^{(H)}$ that has coefficient vector the i_cth column of P, i.e., equal to $P\mathbf{e}_{i_c}$. The set of functions $\{\phi_{i_c}^{(H)}\}$ forms the coarse basis of interest. The parameter H stands for characteristic size of the support of the coarse basis functions.

The above matrix–vector strong approximation property (1.1) admits the following finite element function form:

$$\|A\|\ a(u_h - u_H,\ u_h - u_H) \le C_A \sum_{\mathbf{x}_i \in \mathcal{N}_h} f_i^2 = C_A \sum_{\mathbf{x}_i \in \mathcal{N}_h} \left(\int\limits_{\Omega} f(\mathbf{x}) \varphi_i\ d\mathbf{x} \right)^2.$$

Using Cauchy-Schwarz inequality, we have

$$\sum_{\mathbf{x}_i \in \mathcal{N}_h} \left(\int\limits_{\Omega} f(\mathbf{x}) \varphi_i\ d\mathbf{x} \right)^2 \le \sum_{\mathbf{x}_i \in \mathcal{N}_h} \int\limits_{\text{support }(\varphi_i)} f^2(\mathbf{x})\ d\mathbf{x} \int\limits_{\Omega} \varphi_i^2\ d\mathbf{x}.$$

For a fairly general class of basis functions (including piecewise linears) on a quasiuniform mesh, we have

$$\int\limits_{\Omega} \varphi_i^2\ d\mathbf{x} \simeq |\text{support }(\varphi_i)| \simeq h^d.$$

Due to the bounded overlap of the supports of the finite element basis functions, we also have

$$\sum_{\mathbf{x}_i \in \mathcal{N}_h} \int_{\text{support } (\varphi_i)} f^2(\mathbf{x})\, d\mathbf{x} \leq \kappa\, \|f\|_0^2.$$

Thus, we arrive at the energy error estimate of our main interest (using (1.2))

$$a(u_h - u_H,\ u_h - u_H) \leq C_A \kappa\, \frac{h^d}{\|A\|} \|f\|_0^2 \simeq C_A \kappa\, h^2\, \|f\|_0^2.$$

In practice, we typically have $C_A = \mathcal{O}((\frac{H}{h})^2)$ with a constant in the $\mathcal{O}$ symbol, independent of the two mesh sizes (h and H); see, e.g., Corollary 3.1 later on. Thus, we get the following final upscaling energy error estimate:

$$a(u_h - u_H,\ u_h - u_H) \leq c_A\, H^2 \|f\|_0^2. \tag{1.3}$$

In the remaining sections, we summarize a few AMG methods that provide *strong approximation property*.

2 Efficient window-based spectral AMG methods

In [FVZ05] (see also [Va08]), the following AMG method was proposed that exhibits strong approximation property. The original version tends to lead to relatively large coarse spaces so that the resulting two (and multi)–level methods have unacceptably high complexities. In the present section, we propose several approaches in the attempt to reduce the complexity of the original method.

Given an overlapping partition $\{w\}$ of the set of indices $i = 1, 2, \ldots, n$, we extract the rows of a given $n \times n$ matrix A with indices from any given set (called window) w. The respective rectangular matrix is denoted by A_w. By proper reordering, A_w can be written as follows

$$A_w = [A_{ww},\ A_{w,\chi}]\,.$$

Here, A_{ww} is the principal submatrix of A (row and column indices from w) and $A_{w,\chi}$ is the submatrix of A with columns outside w (and row indices from w).

We are interested, for a proper nonnegative diagonal matrix D_w, in the normal matrices $A_w^T D_w A_w$. The diagonal matrices D_w provide a partition of unity, i.e., if I_w stands for extension by zero outside the set w, then $\sum_w I_w D_w I_w^T = I$. This property ensures that

$$\sum_w \mathbf{v}^T A_w^T D_w A_w \mathbf{v} = \sum_w \mathbf{v}^T A^T I_w D_w I_w^T A \mathbf{v} = \|A\mathbf{v}\|^2, \qquad A_w = I_w^T A. \tag{2.1}$$

The method in question uses the symmetric semi-definite Schur complements S_w defined as follows:

$$\mathbf{v}_w^T S_w \mathbf{v}_w = \inf_{\mathbf{v}_\chi} \begin{bmatrix} \mathbf{v}_w \\ \mathbf{v}_\chi \end{bmatrix}^T A_w^T D_w A_w \begin{bmatrix} \mathbf{v}_w \\ \mathbf{v}_\chi \end{bmatrix}. \tag{2.2}$$

The original method utilizes the eigenvectors of the semidefinite Schur complements S_w,

$$S_w \mathbf{p}_k = \lambda_k \ \mathbf{p}_k, \ k = 1, \ \ldots, \ n_w. \tag{2.3}$$

For efficiency reason, for a given tentative interpolation matrix $\widetilde{P}$, we use in (2.3) instead the modified Schur complements

$$\mathbf{v}_w^T S_w \mathbf{v}_w = \inf_{\mathbf{v}_\chi \in I_\chi^T \mathrm{Range}(\widetilde{P})} \begin{bmatrix} \mathbf{v}_w \\ \mathbf{v}_\chi \end{bmatrix}^T A_w^T D_w A_w \begin{bmatrix} \mathbf{v}_w \\ \mathbf{v}_\chi \end{bmatrix}. \tag{2.4}$$

In what follows, we denote the exact window Schur complement with S_w^*.
We first form local interpolation matrices P_w by putting together the first $n_w^c \geq 1$ eigenvectors (in the lower part of the spectrum of S_w), i.e.,

$$P_w = \left[\mathbf{p}_1, \ \ldots, \ \mathbf{p}_{n_w^c} \right]. \tag{2.5}$$

The corresponding eigenvalues (ordered in an increasing order) are such that $\lambda_k \leq \mathrm{tol}\ \lambda_{\max}(S_w)$ for $k \leq n_w^c$ and

$$\lambda_k(S_w) > \mathrm{tol}\ \|S_w\| = \mathrm{tol}\ \lambda_{\max}(S_w) \ \text{ for } k > n_w^c. \tag{2.6}$$

Here, we have the freedom to choose the pre-selected tolerance "tol" (a number between zero and one) that may also vary with w.
The eigenvectors $\{\mathbf{p}_k\}_{k=1}^{n_w}$ are orthogonal and assumed normalized.
The global P is computed based on another partition of unity set of nonnegative $n_w \times n_w$ diagonal matrices $\{Q_w\}$ that satisfy

$$I = \sum_w I_w Q_w I_w^T.$$

Then P is defined as follows

$$P = \sum_w I_w Q_w \left[0, \ P_w, \ 0\right] = \sum_w I_w Q_w P_w (I_w^c)^T. \tag{2.7}$$

Here, I_w^c maps the local indices of the eigenvectors coming from the window w to their global indices expanding the result with zeros elsewhere. Thus we have defined a process that from a tentative $\widetilde{P}$ produces another one P. This can be iterated several times (by possibly changing the parameters such as $\{w\}$ and tol). In the next theorem, we formulate conditions ensuring that P admits a strong approximation property.

Theorem 2.1 *Consider the iterated window spectral AMG interpolation matrix P constructed on the basis of the modified window Schur complements using a $\widetilde{P}$ that satisfies the following estimate*

$$\sum_w \|D_w^{\frac{1}{2}} A_{w,\,\chi} (\mathbf{v}_\chi - I_\chi^T \widetilde{P} \mathbf{w}_c)\|^2 \leq \mu \ \|A\mathbf{v}\|^2. \tag{2.8}$$

That is, $\widetilde{P}$ is such that for any $\mathbf{v}$, when restricted to a complementary set χ, there is a coarse vector $\mathbf{w}_c$ (depending on $\mathbf{v}$ and the set χ) such that for a fixed number $\mu > 0$ (2.8) holds. Then, if we choose tol $= \frac{1}{\delta} \leq 1$ in the two-level spectral decomposition defining the local P_w so that (see (2.6))

$$\|S_w\| \leq \delta \, \lambda_{m_w+1}(S_w),$$

and if we also assume the quasiuniformity of the windows, i.e., the estimate

$$\beta \, \|A\|^2 \leq \|S_w^*\|, \tag{2.9}$$

then, the following main strong approximation property holds for P

$$\|A\|^2 \|\mathbf{v} - P\mathbf{v}_c\|^2 \leq \eta \, \|A\mathbf{v}\|^2. \tag{2.10}$$

Here, $\eta = \frac{\delta}{\beta} \, (1 + \sqrt{\mu})^2$, where μ is from (2.8).

3 A new "window"-based spectral AMG method for finite element matrices

Here, we present a modified version of the method applied to finite element matrices A. The difference is in the eigenproblems that we use. Also, it utilizes a special partition of unity matrices. A main additional assumption is that the window sets are covered exactly by fine–grid elements and that we have access to the respective fine–grid element matrices so that we can assemble the semi–definite local matrices further denoted by Λ_w. Therefore, we have the identity $A = \sum_w I_w \Lambda_w I_w^T$.

We solve eigenproblems associated with the pair of matrices Λ_w and S_w, where S_w is the exact window-based Schur complement (as introduced before). The eigenproblems read (compare with (2.3)):

$$S_w \mathbf{p}_k = \lambda_k \, \Lambda_w \, \mathbf{p}_k, \; k = 1, \ldots, \; n_w, \tag{3.1}$$

where the eigenvalues are numbered in an increasing order and the eigenvectors are Λ_w–normalized.

Since, the matrices Λ_w can also be only semi–definite, to have real eigenvalues the nullspace of Λ_w should be contained in the nullspace of S_w, which is the case for finite elliptic matrices (Laplacian–like as well as elasticity).

It is clear that we can choose the eigenvectors $\mathbf{p}_k$ be orthogonal to the nullspace of Λ_w (and Λ_w–orthogonal to each other). Let the columns of $P_w^{(0)}$ span the nullspace of Λ_w. Then, we have $\mathbf{p}_k^T P_w^{(0)} = 0$.

Based on a preselected tolerance tol $\in [0, 1)$, we choose n_w^c such that $\lambda_k >$ tol λ_{n_w} for $k > n_w^c$. The local interpolation matrices are defined similarly as before (cf., (2.5)), now augmented with the nullspace, i.e.,

$$P_w = \left[P_w^{(0)}, \; \mathbf{p}_1, \; \ldots, \; \mathbf{p}_{n_w^c} \right]. \tag{3.2}$$

To define the global one, we use special diagonal matrices $\{Q_w\}_w$ with nonnegative entries that provide partition of unity, i.e., we have

$$I = \sum_w I_w Q_w I_w^T.$$

Each Q_w has entries on its diagonal $q_{w,\,i}$, $i \in w$, defined as follows:

$$q_{w,\,i} = \frac{\|\Lambda_w\|}{\sum\limits_{w'\colon\, i\in w'} \|\Lambda_{w'}\|}. \tag{3.3}$$

At the end we formulate our main result.

Theorem 3.1 *Let A be a given finite element s.p.d. matrix. Consider a given set of windows $\{w\}$ where each window w is exactly covered by fine–grid elements. Assume also that the local finite element matrices Λ_w corresponding to the sets w are available. The nullspace (if nonempty) of the local matrices Λ_w is assumed known (explicitly computed). That is, let the nullspace of Λ_w be represented by the range of an explicitly available local matrix $P_w^{(0)}$. Assume that this nullspace is contained in the nullspace of the window Schur complement S_w (defined in* (2.2)*). The global interpolation matrix P is defined as in* (2.7) *based on the local interpolation matrices* (3.2) *and the weights $q_{w,\,i}$ (entries of the diagonal partition of unity matrices Q_w) are defined in* (3.3)*. Then, the following global strong approximation property holds*

$$\|\mathbf{v} - P\mathbf{v}_c\|_A^2 \le \kappa \ \max_w \ Cond^+(\Lambda_w) \ \max_w \ \left(\frac{1}{tol\ \lambda_{\max}(\Lambda_w^+ S_w)}\right) \ \|A\mathbf{v}\|^2.$$

Above, $\kappa \ge 1$ depends on the overlap of $\{w\}$, $tol \in (0,1]$ (in general depending on w) is the tolerance used to define the portion of the eigenvectors $\mathbf{p}_k$ in the lower part of the spectrum computed in (3.1) *used to define P_w, $Cond^+(\Lambda_w) = \frac{\|\Lambda_w\|}{\lambda^+_{\min}(\Lambda_w)}$ is the effective condition number of Λ_w computed in a subspace orthogonal to the nullspace of Λ_w. Finally, $\lambda_{\max}(\Lambda_w^+ S_w) = \max_k \lambda_k$ where λ_k are from* (3.1)*.*

Corollary 3.1 *For finite element s.p.d. matrices A coming from second order elliptic problems, the constructed finite element modification of the window-based spectral AMG method, the following strong approximation property holds*

$$\|\mathbf{v} - P\mathbf{v}_c\|_A^2 \le C\ \kappa\ \left(\frac{H}{h}\right)^2 \ \max_w \ \frac{1}{tol\ \|A_{w_0,w_0}\|}\ \|A\mathbf{v}\|^2.$$

References

[FVZ05] R. Falgout, P. S. Vassilevski, and L. T. Zikatanov, "*On two–grid convergence estimates,*" Numerical Linear Algebra with Applications, **12**(2005), pp. 471–494.

[Va08] Panayot S. Vassilevski, "Multilevel Block Factorization Preconditioners, Matrix-based Analysis and Algorithms for Solving Finite Element Equations," Springer, New York, 2008. 514 p.